Using intuition or a formal palliative care needs assessment screening process in general practice to predict death within 12 months

Palliative Medicine

Using intuition or a formal palliative care needs assessment screening process in general practice to predict death within twelve months: a randomised controlled trial.

Journal:	*Palliative Medicine*
Manuscript ID	PMJ-16-0213.R2
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Mitchell, Geoffrey; University of Queensland, Discipline of General Practice; Senior, Hugh; Massey University, College of Health; University of Queensland Faculty of Medicine and Biomedical Sciences, General Practice Rhee, Joel; University of New South Wales, School of Public Health and Community Medicine Ware, Robert; Griffith University, Menzies Health Institute Queensland, Young, Sharleen; West Moreton Hospital and Health Service; University of Queensland, Discipline of General Practice Teo, Patrick ; University of New South Wales, School of Public Health and Community Medicine Murray, Scott; University of Edinburgh, General Practice - Research Boyd, Kirsty; University of Edinburgh, General Practice Clayton, Josephine; Greenwich Hospital, Palliative and Supportive Care Service; University of Sydney, Sydney Medical School
Keywords:	general practice, patient screening, end of life, frailty, multimorbidity, care planning
Abstract:	Background: Population ageing will lead to more deaths with an uncertain trajectory. Identifying patients at risk of dying could facilitate more effective care planning. Aim: To determine whether screening for likely death within twelve months is more effective using screening tools or intuition. Design: RCT of screening tools (ST) (Surprise Question (SQ) plus the Supportive and Palliative Care Indicators Tool (SPICT) for SQ+ patients) to predict those at risk of death at 12 months compared with unguided intuition (I). Clinical trials registry ACTRN12613000266763. Setting/Participants: Australian general practice. Thirty GPs (ST-12, I-18) screened all patients (n=4365) aged ≥70 years seen at least once in the last two years. Results: There were 142 deaths (ST 3.1%, I 3.3%: p=0.79). GPs identified more at risk of dying using SQ (11.8%) than intuition (5.4%: p=0.01), but no difference with SQ+ then SPICT (5.1%: p=0.87). SQ+ predicted more deaths (53.2%; I 33.7% p=0.001), but SQ+/SPICT predictions were similar (5.1%, p=0.87 vs intuition). There was no difference in proportions correctly predicted to die (SQ 1.6%; I 1.1% p=0.156, SQ+/SPICT 1.1%;

	p= 0.86 vs intuition). ST had higher sensitivity and lower specificity than intuition, but no difference in positive or negative predictive value. Conclusions: ST was better at predicting actual death than intuition, but with a higher false positive rate. Both were similarly effective at screening the whole cohort for death. Screening for possible death is not the best option for initiating end of life planning: recognising increased burden of illness might be a better trigger.

Using intuition or a formal palliative care needs assessment screening process in general practice to predict death within twelve months: a randomised controlled trial.

Geoffrey K Mitchell MBBS, PhD[1] (Corresponding author)

1. School of Medicine, University of Queensland

Herston, 4029, Queensland, Australia

Email: g.mitchell@uq.edu.au

Hugh E Senior DPH, PhD[1,2]

2. College of Health, Massey University, Auckland, New Zealand

Joel J Rhee MBBS(Hons) PhD[3,4]

3. HammondCare Centre for Positive Ageing and Care, Sydney.
4. School of Public Health and Community Medicine,
University of New South Wales.

Address: School of Public Health and Community Medicine
The University of New South Wales
UNSW SYDNEY, 2052 NSW, Australia

Robert S Ware, PhD[5]

5. Menzies Health Institute Queensland, Griffith University
Address: 170 Kessels Road, Nathan, 4111,QLD, Australia
Email: r.ware@griffith.edu.au

Sharleen Young[1,6]

6. West Moreton Hospital and Health Service,

Address: Ipswich Hospital, Ipswich, 4305, Queensland, Australia.

Patrick CK Teo BMed, MD [4]

Scott Murray MD, FRCGP[7]

7. Primary Palliative Care Research Group, The Usher Institute for Population Health Sciences & Informatics, The University of Edinburgh Medical School, Teviot Place, Edinburgh, EH8 9AG, Scotland, UK

Kirsty Boyd MBChB, FRCP[7]

Josephine M Clayton FRACP, PhD[8,9]

8. HammondCare Palliative and Supportive Care Service, Greenwich Hospital, Greenwich, NSW, Australia

9. Northern Clinical School, School of Medicine, Sydney, NSW, Australia

Acknowledgements

We thank the GPs and practice staff of the practices involved in this study. We acknowledge the role of the Palliative Care Scientific Working Group of the Primary Care Collaborative Cancer Clinical Trials Group (PC4) in the early development of this concept, a clinical trials group funded by Cancer Australia. We also acknowledge Dr Jian (Aaron) Tan who assisted in data collection and analysis. We thank doctors and staff of Limestone Medical Centre, Ipswich, Australia, where the pilot work was undertaken.

Funding statement: The researchers gratefully acknowledge the RACGP Foundation for their support of this project. This study was funded by the Primary Care Collaborative Cancer Clinical Trials Group (PC4)/RACGP Research Grant (2012-3).

Competing interest statement: The authors declare no completing interests.

Access to the data: The dataset from this trial can be obtained by contacting the first author.

Word count:

Abstract 249

Text – 3285 (excluding tables and references)

Abstract

Background: Population ageing will lead to more deaths with an uncertain trajectory. Identifying patients at risk of dying could facilitate more effective care planning.

Aim: To determine whether screening for likely death within twelve months is more effective using screening tools or intuition.

Design: RCT of screening tools (ST) (Surprise Question (SQ) plus the Supportive and Palliative Care Indicators Tool (SPICT) for SQ^+ patients) to predict those at risk of death at 12 months compared with unguided intuition (I). Clinical trials registry ACTRN12613000266763.

Setting/Participants: Australian general practice. Thirty GPs (ST-12, I-18) screened all patients (n=4365) aged ≥70 years seen at least once in the last two years.

Results: There were 142 deaths (ST 3.1%, I 3.3%: p=0.79). GPs identified more at risk of dying using SQ (11.8%) than intuition (5.4%: p=0.01), but no difference with SQ^+ then SPICT (5.1%: p=0.87). SQ^+ predicted more deaths (53.2%; I 33.7% p=0.001), but SQ^+/SPICT predictions were similar (5.1%, p=0.87 vs intuition). There was no difference in proportions correctly predicted to die (SQ 1.6%; I 1.1% p=0.156, SQ^+/SPICT 1.1%; p= 0.86 vs intuition). ST had higher sensitivity and lower specificity than intuition, but no difference in positive or negative predictive value.

Conclusions: ST was better at predicting actual death than intuition, but with a higher false positive rate. Both were similarly effective at screening the whole cohort for

death. Screening for possible death is not the best option for initiating end of life planning: recognising increased burden of illness might be a better trigger.

Keywords: general practice, patient screening, end of life, frailty, mulltimorbidity, care planning

Key statements

What is already known about the topic?

- Care planning for the end of life improves outcomes.
- Identifying patients for whom end of life care planning will benefit is difficult.

Several tools have been produced to help identify those approaching the end of life and whose care needs will escalate. It is not known how accurate these are at predicting death.

What this paper adds?

- Using a combination of screening tools to screen patients 70 years and older in general practice lists was able to identify patients at risk of dying in 12 months better than unguided intuition.
- However, screening GP lists of patients age 70 and older, using these tools is no more accurate than intuition in predicting death at 12 months.
- Both intuition and the screening process we tested are several times better at predicting death than the incidence rates of death in the intuition and screening tool populations, but have unacceptably high levels of false positive results for predicting dying within 12 months.

Implications for practice, theory or policy?

- Using current screening tools to screen general practice patient lists to predict dying within 12 months is not feasible.

- Different triggers to generating care plans in the later stages of life are required. Sentinel events like unexpected hospitalisation may be better markers for the need for care planning.

Background

Countries worldwide face ageing populations. As people age, their burden of illness rises. Most will die after living for many years with conditions such as frailty, multimorbidity, dementia and organ failure, and relatively fewer will die of cancer.[1]

By contrast, the patients of specialist palliative care services suffer from cancer. [2] In Western Australia, only 8% of people who died of a non-malignant disease accessed specialised palliative care services, while 68% of people who died of cancer did so.[3] The trajectory of decline in non-cancer disease is more uncertain than that of cancer, and the trajectory to death can be much longer.[4] Planning end of life care for these conditions is challenging.

The last period of an individual's life incurs a high proportion of lifetime health expenditure.[5] In the USA, different functional trajectories of illness varied in speed and intensity of deterioration in the year before entering a hospice program.[6] Forward planning for anticipated deterioration may lessen its impact, may allow a person to remain at home, may deliver a better sense of control over the situation and more peace of mind.

An approach to this problem has been to identify patients at risk of dying within a foreseeable timeframe, usually several months, to then generate an end of life care plan that can be enacted when deterioration to death occurs. [7]

Numerous instruments aimed at identifying those people at risk of deteriorating towards death by recognising the presence of particular physical or other markers.[8] The simplest of these is the Surprise Question (SQ), which asks practitioners whether they would be surprised if this person were to die within 6-12 months.[9] If the answer is "No, I would not be surprised" this indicates a perceived risk of deterioration to death (referred to as SQ^+ throughout this paper). Other tools developed require practitioners to identify general indicators of physical deterioration, and/or specific illness-based indicators of decline for the person being considered. The utility of this latter approach in identifying individuals at risk of dying has been demonstrated in a range of settings where death in the foreseeable future has high prevalence (For example, in acute hospital inpatients[10]).

General practices in the UK receive financial incentives to maintain palliative care patient registers.[11] Formal screening of general practice patient lists takes time. Many health systems, such as the Australian general practice system do not require or fund system-level planning like this, so identification of at-risk patients needs to be done opportunistically, with the risk that patients will 'slip through the net'. If done systematically, screening needs to be efficient and cost-effective.

We decided to test whether it is it feasible and effective to apply a screening procedure to identify people at risk of dying in the foreseeable future, in the Australian general practice setting, by conducting a randomised controlled trial (RCT) to test whether a

formal screening tool was more effective in predicting foreseeable death than a purely intuitive approach.

Methods

Study type

We undertook a randomised controlled trial of GPs predicting dying in their patients using either intuition or a process involving screening tools. Because this has not been done before, we considered this to be an exploratory RCT, with an aim of the study to be to calculate statistics to inform sample size calculations for future studies.

Primary and other outcomes

The primary outcome was the proportion of patients correctly identified as being at risk of death at twelve months. Secondary outcomes included test characteristics including sensitivity, specificity, positive and negative predictive values, likelihood ratios and pre- and post-test likelihoods of dying or not dying.

Setting and recruitment

Practices were recruited through practice-based research networks of the University of New South Wales and the University of Queensland, Australia; notices in doctor newsletters and direct approaches to practices.

Australian general practice occurs almost exclusively in private clinics, with multiple general practitioners (GPs) and often a practice nurse. There is ready availability of pathology, radiology and community specialist support in most GP settings. Funding is through a universal health insurance scheme which is based on fee for service: GPs rely on regular patient throughput to generate an income stream.

Participant Eligibility

GPs within the participating practices were excluded if they had worked <12 months in the practice; were not seeing patients ≥ 70 years of age; or where the practices did not use computerised patient records.

Randomization procedure

Eligible GPs were randomly assigned to either an intuition (I) group or a screening tool (ST) group with a 1:1 allocation by a computer generated number sequence with stratified permuted blocks of random sizes. GP participants were stratified according to years of general practice experience (≤10 or >10 years).

Screening instruments

We selected the Supportive and Palliative care Indicator Tool (SPICT, 2012 version)[12] which helps clinicians identify people who have supportive or palliative care needs. The SPICT is a two-step clinical guidance tool. The first step seeks general indicators of deteriorating health, the second seeks specific indicators of advanced disease for a range of specific conditions including cancer, dementia/frailty, and system-based diseases. A combination of two general indicators of deterioration and one specific indicator prompts assessment of supportive or palliative care needs (SPICT^{+}). The SPICT has been tested in a range of conditions in hospital[10], but not as a screening tool in general practice.

Our pilot work indicated that the SPICT took some minutes to complete for each patient, so on its own would have been impractical as a screening tool in the study setting. Hence we integrated the SQ as a pre-screening tool prior to the full SPICT. We also identified that the vast majority of deaths in primary care occurred in patients ≥70 years.

Study Procedure

After providing informed consent, we generated a patient list comprising individuals aged ≥70 years old who had been seen by study GPs in the last two years. Patients not correctly allocated to their lists were reassigned to the correct GP. GPs were asked to identify people at risk of dying using one of two methods.

One group (Intuition) was asked to search their list to identify patients they thought might die within 12 months, with no external prompts or guides. The actual question asked was: *"Please complete one form for each patient under your care and whom you have added to your list of patients who will most likely die in the next 12 months."*

The second group (ST) initially screened their patient lists using the SQ, and then screened those using the SPICT indicators to identify patients who might die. The SQ responses could be SQ^{+} (answered no to the SQ), SQ^{-} (answered yes to the SQ) or uncertain about the possibility of death within 12 months (SQ^{u}). The GP then completed

the SPICT on SQ^{+} and SQ^{U} patients. Patients who were SQ^{-} were not considered further. GPs from both groups were also asked to add any patients aged <70 whom they considered at risk of death, to their lists. Intuition group GPs were asked to describe briefly the reasons for selecting the patients considered at risk of dying.

Blinding

Each practice contained GPs in both allocation groups. In order to minimize contamination, those randomised to intuition were given their instructions first. Once they had completed screening, the ST GPs in that practice were given their instructions and commenced screening.

Follow-up

We sought deaths that occurred up to twelve months post-baseline through the death registries of Queensland and New South Wales, searched eighteen months after baseline to maximize the capture of these deaths.

Analysis

Categorical variables were summarized as frequency (percentage). Participant GP characteristics within the groups were assessed using Fisher's Exact Test. Because of the different ways patient characteristics were assessed, a descriptive comparison was made, as a statistical comparison was not possible. To calculate diagnostic statistics for intuition and ST groups in predicting death within twelve months we used a generalized

linear model with binomial family, identity link, and robust standard errors to adjust for possible correlation with GPs. Effect estimates are reported as mean difference (MD) and 95% confidence interval (95%CI) Likelihood ratios and post-test probabilities were calculated using the overall probability of death as the pre-test probability. We calculated the intraclass correlation coefficient (ICC) to inform future power calculations for studies of this type. [12] Significance was set at $P < 0.05$.

Qualitative data collection and analysis

Semi-structured interviews were conducted with participants to determine the feasibility and acceptability of adopting a case finding approach in general practice/primary care; and to explore the barriers and facilitators to implementing this. Participants were selected to ensure that views of both male and female participants, different ages and clinical experience, practice types and geographical locations were represented. The interviews were conducted by telephone, digitally recorded then transcribed. Qualitative descriptive methodology informed the analysis. All the coding and analysis was undertaken by one investigator (JR); in addition HS and PT jointly coded one-third of the interviews and had coding and analysis meetings with JR in order to increase the data's trustworthiness.

Ethics and Trial Registration

The project was approved by Ethics committees of the University of Queensland (2012001275) and the University of New South Wales (HC12553). The trial was

registered with the Australian and New Zealand Clinical trials register, number ACTRN12613000266763.

Results

General Practitioners and patients

We recruited 40 GPs from nineteen practices from November 2012 to August 2013. One GP dropped out after recruitment but before randomization. Nineteen were allocated to the screening tools (ST) process and twenty to the intuition (I) process. We retrieved data from eighteen of nineteen (95%) intuition group GPs but only twelve of 19 (63%) ST GPs, most citing pressures of work for failure to provide data. (Figure 1). There was no difference in the proportions of age or years in practice between the GP groups, but more practitioners in the Intuition group trained in countries other than Australia. The difference in gender balance between the groups approached significance (Table 1) Patient characteristics appeared similar in each group. (Table 2) The median number of patients over 70 assessed by intuition GPs was 129 (range 11, 321; Interquartile range (IQR) 73-180.25) and ST was 94 (range 8-481, IQR 52-224.25). The percentage of patients where death was predicted ranged from 0.0% to 29.1%, with median = 6.6% , when the SPICT positive question was included these figures fell to median = 3.4%, range = 0.0% to 29.1%).

ST GPs assessed 1525 patients, and intuition GPs, 2840 patients. Similar proportions of patients from each group died (Table 3). Because GPs within the one practice allocated to different groups saw the same patient, these patients were counted in the assessments more than once.

Patients identified as at risk of dying (Primary outcome)

When ST GPs initially applied the SQ, they identified 179 of 1525 patients (11.7%) who might die within twelve months compared with 154 of 2840 (5.4%) identified by intuition group GPs, giving a mean difference (MD) between groups of 6.3% (95%CI: 1.4%, 11.2%; P=0.01). Eight of the identified patients were aged under 70 years (ST=2, I=6). When the SPICT was conducted on these 179 patients initially identified with SQ, the number of SPICT^{+} patients (i.e those continuing to be identified as being at risk of death) fell to 78 (5.1%), a similar proportion as for intuition alone (MD=-0.3%; 95%CI: -4.0%, 3.4%; P=0.87)

Deaths recorded at twelve months

The death registries recorded 142 deaths, 95 (3.3%) in the intuition group and 47 (3.1%) in the ST group (MD = -0.3%; 95%CI: -2.2%, 1.6%, P=079). There were no deaths for those people under the age of 70. (Table 3)

For patients who died, GPs in the ST group predicted more deaths (25/47; 53.2%) than those in the intuition group (32/95; 33.7%), MD=19.5%;95%CI 7.6%, 31.4%; P=0.001). However, for all patients reviewed by ST and intervention GPs, the proportion of deaths was similar (ST 25/1525 (1.6%); I 32/2840 (1.1%) MD=0.5%; 95%CI: -0.5%, 1.5%; P=0.33). When the SQ+ then SPICT process was applied, the proportion of identified deaths fell to 16/1525 (1.1%) MD = 0.0%; 95%CI: -0.9%, 0.8%; P=0.86).

Comparative ability of intuition and formal tools to predict death

Sensitivity and specificity of intuition and predictive tools were significantly different, with a prediction of SQ^{+} plus SQ^{U} being more sensitive and less specific than intuition. The positive and negative predictive values of both groups were similar (Table 4).

While screening with SQ then SPICT reduced the numbers considered at risk, fewer deaths were correctly identified. There was no difference in sensitivity, specificity, positive or negative predictive value compared with intuition alone (Table 4).

The odds of predicting dying are related to the likelihood ratios of a positive and negative test and the prevalence of dying at 12 months in each group. The odds of dying with a positive test, or not dying with a negative test, are very similar when either intuition, SQ alone, or SQ/SPICT process were applied (Table 4). In each case, the rate of detection was between five to seven times the actual death rate of the population. (Table 5)

Intracluster correlation coefficient (ICC) estimation derived from this study

To inform future trials of this type we calculated ICCs to measure the extent to which there are systematic differences between GPs when predicting death in their patients. When considering the Surprise Question alone, ICC = 0.047, and when the screening question was combined with a SPICT positive, ICC = 0.049. If a future study had the

same median number of patients per GP as this study (n=124), then the design effect of the trial would be 6.8.

Qualitative findings

Eleven GP participants (5 female, 6 male) were interviewed. Participation in the study led them to consider end of life care management in their routine clinical practice positively. They felt that limited time and the opportunity cost of screening were important barriers to routine uptake of end of life screening. However they also identified that incorporating end of life screening tools into electronic medical records would assist in raising awareness of possible deterioration towards death, though the use of automated prompts on computer records when certain criteria were present, or prompting consideration of the SQ and SPICT in routine health assessments for older people. (Table 6)

Discussion

Predicting death in general practice is difficult. Screening tools allowed better prediction of death than relying on intuition alone, but only within the people considered at risk. For the entire screened population, screening tools were no better than intuition in predicting death. However, attempting to identify people at risk of dying by any means raised the recognition of impending death approximately five- to seven-fold. All of the processes tested suffered from low sensitivity and high false positive rates.

The high false positive rate is due to the low prevalence of dying in general practice. The higher the prevalence of a condition, the better the test characteristics. For example, Moroni and colleagues asked the GPs of patients with advanced cancer to apply the SQ to them, and demonstrated a positive predictive value of 83.8% at 12 months post assessment [13] compared with our result of 19%.

The objective of identifying when someone enters the last year of life is to ensure that they and their family can be afforded the least distress and greatest comfort by providing the best possible care. Accurate prediction of a person's death is not the main objective. Responses to patients identified by screening as nearing the end of life have to be timely and feasible for general practices, to be effective for patients and their carers.

Active care planning in the final months of life improves quality of life[14] , reduces

hospitalization[15,16] and maintains function.[15]. However, in the Australian service environment where time is at a premium, the GP response to patients identified by routine screening would almost certainly be inadequate, as large numbers of false positive patients would all require a response.

Therefore, accurate prediction of death is probably not the best signal of the need for end-of-life care planning in general practice. Recognising that the burden of illness is increasing may be a more appropriate goal. Intensive clinical care planning and service provision for all persons identified is not feasible, rather escalating levels of input by the primary care team as the person's needs increase would be the ideal response. This approach could use a significant event like hospitalization or the onset of a new medical condition to trigger an assessment of the risk of dying.[17] Practice computers could be used to flag such events automatically, but the most important research task is to identify what these events are. [18]

Strengths and weaknesses of the study

We used an RCT to produce high-level evidence by ensuring balance in GP and patient characteristics prior to implementing the interventions. We believe the randomisation and approaches taken to minimise contamination taken were robust. However, this is a pilot study and the results have to be treated with caution. A major strength of the study was the use of state death registries to identify deaths.

There are weaknesses to the study. The use of intuition was a proxy for normal care, but asking the GPs to review lists and use intuition alone is in fact an intervention, and probably does not reflect normal practice. (Australian GPs are not encouraged to maintain palliative care registers) It can also be argued that using the SQ to identify patients at risk of dying is also purely intuitive, and so the trial was in effect testing only the ability of the SPICT. However, the proportion of people identified using the SQ was over twice that of asking GPs to use intuition alone, without the "surprise" descriptor. The SQ clearly guides intuition and does it very effectively. This study required GPs to report on any patient they had seen on the list over 70 in the last two years. They may have misclassified patients for whom their knowledge was limited. Further, as the GPs manually checked the lists, patients may have been inadvertently overlooked. We experienced differential drop out rates, with more GPs in the SQ/SPICT group dropping out compared to the intuition group. This highlights the resource intensive nature of reviewing the list of patients systematically using prediction tools.

Since all of the practices were group practices, more than one GP sometime assessed the same patient. As this was a pragmatic trial and not a population-based study of death rates, duplicate observations by multiple GPs of some patients was expected and not a study weakness. The number of GPs not trained in Australia was significant, but numbers are small and it is unlikely to be of practical importance. The difference in proportion of female GPs in each group approached significance. Since these numbers

are of randomized GPs before dropouts, the difference was not likely to introduce systematic bias into the trial.

Conclusion

Screening for critical, low prevalence conditions normally requires a simple, easily administered test that has a low false positive rate. GP response to the test should prevent adverse consequences like preventable emergencies. Accurate prognostication of dying several months from the event is very difficult in the general practice setting due to the low prevalence of the index event. It may be better to screen general practice patients for risk of accelerating deterioration towards death, rather than for the risk of death itself within a specified time frame. Reliable signals of this deterioration in general practice are yet to be determined.

Health care systems that do not embed population-based surveillance in general practice, such as that in Australia, require novel approaches to identify and then deliver the best care for people who are approaching the end of their lives. Further research is needed to determine the best ways to identify the people in need of supportive care efficiently, so that the required planning and care is provided as effectively as possible.

References

1. Gill TM, Gahbauer EA, Han L, Allore HG. Trajectories of disability in the last year of life. N Engl J Med. 2010;362:1173-80.

2. Palliative Care Australia. Palliative care service provision in Australia: a planning guide. Canberra: Palliative Care Australia; 2003.

3. Rosenwax LK, McNamara BA. Who receives specialist palliative care in Western Australia--and who misses out. Palliat Med. 2006;20:439-45.

4. Lynn J. Perspectives on care at the close of life. Serving patients who may die soon and their families: the role of hospice and other services. JAMA. 2001;285:925-32.

5. Alemayehu B, Warner K. The Lifetime Distribution of Health Care Costs. Health Serv Res. 2004;39:627-42.

6. Stabenau HF, Morrison LJ, Gahbauer EA, Leo-Summers L, Allore HG, Gill TM. Functional trajectories in the year before hospice. Ann Family Med. 2015;13:33-40.

7. Thomas K. Caring for the Dying at Home: Companions on the Journey Abdingdon: Radcliffe; 2003.

8. Maas EA, Murray SA, Engels Y, Campbell C. What tools are available to identify patients with palliative care needs in primary care: a systematic literature review and survey of European practice. BMJ Support Palliat Care. 2013;3:444-51.

9. Pattison M, Romer AL. Improving Care Through the End of Life: launching a primary care clinic-based program. J Palliat Med. 2001;4:249-54.

10. Highet G, Crawford D, Murray SA, Boyd K. Development and evaluation of the Supportive and Palliative Care Indicators Tool (SPICT): a mixed-methods study. BMJ Support Palliat Care. 2014;4:285-90.

11. BMA, NHS Employers. Quality and Outcomes Framework for 2012/13. Guidance for PCOs and practices. London: BMA, NHS Employers, 2012.

12. Donner A, Klar N. Design and analysis of cluster randomization trials in health research. London: Arnold; 2000.

13. Moroni M, Zocchi D, Bolognesi D, Abernethy A, Rondelli R, Savorani G, et al. The 'surprise' question in advanced cancer patients: A prospective study among general practitioners. Palliat Med. 2014;28:959-64.

14. Mitchell GK, Del Mar CB, O'Rourke PK, Clavarino AM. Do case conferences between general practitioners and specialist palliative care services improve quality of life? A randomised controlled trial (ISRCTN 52269003). Palliat Med. 2008;22:904-12.

15. Abernethy AP CD, Shelby-James T, Rowett D, May F, Samsa GP, Hunt R, Williams H, Esterman A, Phillips PA. . Delivery Strategies to Optimize Resource Utilization and Performance Status for Patients Advanced Life-Limiting Illness: Results From the ''Palliative Care Trial'' [ISRCTN81117481]. J Pain Symptom Manage. 2013;45:488-505.

16. Mitchell G, Zhang J, Burridge L, Senior H, Miller E, Young S, et al. Case conferences between general practitioners and specialist teams to plan end of life care of people with end stage heart failure and lung disease: an exploratory pilot study. BMC Palliat Care. 2014;13:24.

17. Clark D, Armstrong M, Allan A, Graham F, Carnon A, Isles C. Imminence of death among hospital inpatients: Prevalent cohort study. Palliat Med. 2014;28:474-9.

18. Mason B, Boyd K, Murray SA, Steyn J, Cormie P, Kendall M, et al. Developing a computerised search to help UK General Practices identify more patients for palliative care planning: a feasibility study. BMC Fam Pract. 2015;16:99.

Figure 1 – Participant (GP) Flow chart

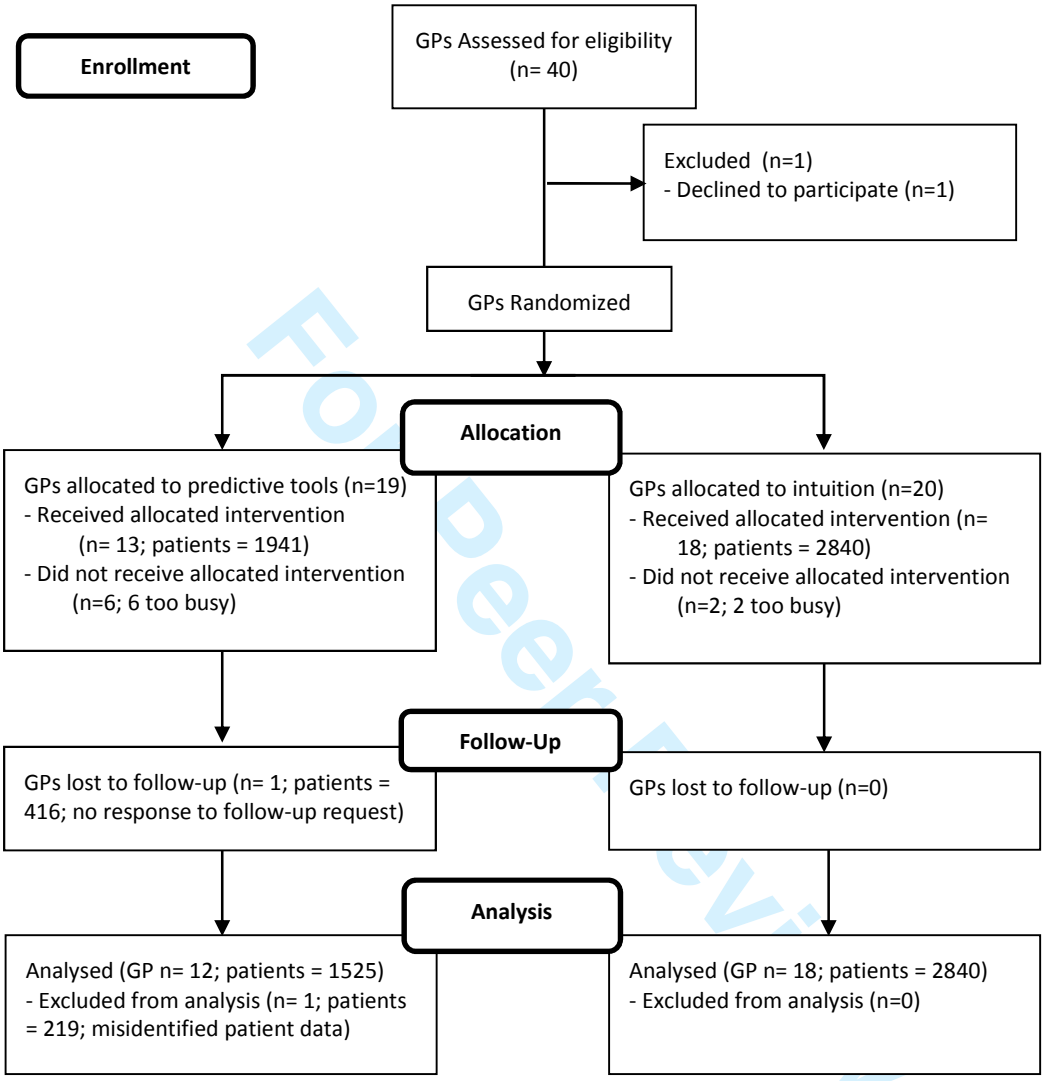

Table 1 GP Characteristics (n=39)

	Intuition (n=20)	Screening Tools (n=19)	P
Sex			
Female	6 (40%)	11 (58%)	0.08
Age[1]			1.00
<=40	6 (35%)	7 (41%)	
41-50	3 (18%)	3 (18%)	
51-60	6 (35%)	6 (36%)	
61+	2 (12%)	1 (6%)	
Country of training			0.34
Australia/ New Zealand	15 (79%)	17 (94%)	
UK/ Ireland/Other	4 (21%)	1 (6%)	
Years in general practice			0.62
<=10	7 (39%)	8 (44%)	
11 to 20	4 (22%	1 (6%)	
21 to 30	4 (22%)	5 (28%)	
31 to 40+	3 (17%)	4 (22%)	

1. Missing items for intuition group: age, n=3; country of training, n=1; years in practice, n=2. Missing items for screening tools group: age, n=2; country of training, n=1; years in practice, n=1.

Table 2 Characteristics of patients and those predicted to die by general practitioners (GPs).

(Intuition GPs were not given any prompts to identify patients at risk of dying, but were asked to provide reasons for their decision to identify individuals. Therefore, the entries in each column are not directly comparable and no statistical tests were conducted. Individuals could have more than one reason for being at risk of dying)

	Tested by SQ/SPICT (n=1525)	Tested by Intuition (n= 2840)
Sex - Female (n (%))	512 (33.6)	1300 (29.8)
Age - Years (Mean, (SD))	79.1 (6.9)	77.9 (6.3)
	(n=179)	**(n=154)**
Conditions *(SPICT categories in italics* Conditions identified by intuition only in normal type.*)*		
Cardiovascular		
Congestive Heart Failure	21 (11.7%)	35 (22.7%)
Peripheral Vascular Disease	2 (1.1%)	19 (12.3%)
Respiratory		
Respiratory impairment (Severe)	10 (5.6%)	12 (7.8%)
Oxygen therapy	2 (1.1%)	
Respiratory impairment (moderate)	-	23 (14.9%)
Neurological		
Indicators of dementia/frailty:		-
Deterioration	42 (23.5%)	
Dysphasia/Dysphagia	14 (7.8%)	-
Pneumonia/ respiratory infection	1 (0.6%)	-
Dementia	-	34 (22.1%)
Stroke (severe)	-	3 (1.9%)
Other	-	37 (24%)
Endocrine		
Diabetes	-	30 (19.5%)
Diabetes with end organ disease	-	9 (5.8%)
Other endocrine	-	4 (2.6%)
Renal		
CKD stage 4-5	-	15 (9.7%)
(+ with deteriorating health)	6 (3.4%)	
Failure from non-renal cause	1 (0.6%)	
Stopping dialysis	0 (0.0%)	

CKD stage 1-3	-	16 (10.4%)
Gastrointestinal		
Severe Liver disease	0 (0.0%)	4 (2.6%)
Liver transplant indicated	0 (0.0%)	-
Active Peptic Ulcer	-	2 (1.3%)
Cancer		
Functional ability deteriorating from progressive metastatic cancer	19 (6.3%)	
Too frail for oncology treatment or treatment for symptom control	8 (2.3%)	
Cancer localised	-	29 (18.8%)
Cancer disseminated	-	11 (7.1%)
Lymphoma	-	2 (1.3%)
Other		
AIDS	-	1 (0.6%)
Rheumatologic conditions	-	6 (3.9%)

Table 3- Number of Deaths and Proportions of Deaths at 12 Months

Proportion of total population	Intuition (n=2840)	Screening tools (n=1525)	MD[1] (95%CI)	P
Total deaths	95 (3.3%)	47 (3.1%)	-0.3% (-2.2%, 1.6%)	0.79
Predicted deaths	32 (1.1%)	25 (1.6%)	0.5% (-0.5%, 1.5%)	0.33

[1]MD- Mean Difference [2]SQ^+ - Surprise Question positive, SQ^- - Surprise Question negative, SQ^U - Surprise Question uncertain, [3]$SPICT^+$ positive identification by SPICT, $SPICT^-$ negative identification by SPICT,

Table 4 – Test characteristics of screening methods compared to intuition. Differences between groups are presented as mean difference (MD) and 95% confidence interval (95%CI). There were 2840 patients in intuition group (95 deaths) and 1525 patients in screening (+/-SPICT) group (47 deaths)

	Intuition (N=2840) (%)	Screening (N=1525) (%)	Screening vs intuition MD(95%CI)	Screening vs intuition P-value	Screening + SPICT (N=1525) (%)	Screening+ SPICT vs intuition MD(95%CI)	Screening +SPICT vs intuition P-value
Deaths, n(%)	95 (3.3%)	47 (3.1%)	-0.3% (-2.2%, 1.6%)	0.79	47 (3.1%)	-0.3% (-2.2%, 1.6%)	
Predicted deaths, n(%)	154 (5.4%)	179 (11.7%)	6.3% (1.4%, 11.2%)	0.01	78 (5.1%)	-0.3% (-4.0%, 3.4%)	0.87
Sensitivity, % (95%CI)	33.7% (23.1%, 44.2%)	53.2% (48.1%, 58.3%)	19.5% (7.6%, 31.4%)	0.001	34.0% (25.3%, 42.8%)	0.4% (-13.7%, 14.4%)	
Specificity, % (95%CI)	95.6% (93.8%, 97.3%)	89.6% (85.5%, 93.7%)	-6.0% (-10.4%, -1.5%)	0.009	95.8% (93.0%, 98.6%)	0.2 (-3.1%, 3.6%)	
PPV, % (95%CI)	20.8% (12.7%, 28.7%)	14.0% (8.8%, 19.1%)	-6.8% (-16.2,2.7%)	0.16	20.5% (12.6, 28.4%)	-0.3% (-11.5, 10.9%)	0.96
NPV, % (95%CI)	97.7% (96.8, 98.5%)	98.4% (97.5%, 99.2%)	0.7% (-0.5%,1.9%)	0.24	97.9% (96.8%, 99.0	0.2 (-1.2%, 1.6%)	0.78

Table 5 - Likelihood ratios and pre-post-test probability of Intuition, Surprise Question (SQ) screening only and SQ then SPICT

	Intuition	SQ[1] screening	SQ then SPICT[2]
Prevalence of death at 12 months (Overall pre-test probability)	3.3%	3.3%	3.3%
Test positive for possible death at 12 months			
Likelihood ratio (95%CI)	7.6 (5.4-10.6)	5.1 (3.8, 6.9)	8.1 (5.1-12.9)
Post-test probability of dying (95%CI)	20.5% (15.7%, 26.5%)	14.8% (11.4%, 19.1%)	20.5% (15.7%, 26.5%)
Test negative for possible death at 12 months			
Likelihood ratio (95%CI)	0.7 (0.6-0.8)	0.5 (0.4, 0.7)	0.7 (0.6, 0.9)
Post-test probability of dying (95%CI)	2.3% (2.0%, 2.7%)	1.8% (1.3%, 2.4%)	2.3% (1.9, 2.8%)

[1] SQ- Surprise question [2]SPICT- Supportive and Palliative care Indicator Tool

Table 6. General practitioner views on the impact of, and facilitators and barriers to, systematic screening of people for risk of dying.

Impact of study participation on perceptions of end of life care

"I guess it would prompt me to be more proactive about future planning… 'cause at the moment, I probably would be waiting for quite – something quite disastrous to happen to the patient before I brought up anything about care or – they'd have to fall into a heap before that whole topic came up. So, yeah, so I think it would definitely change my management." (N07)

Time and income foregone as barriers to implementation

"It's process that takes time and time is at a premium in general practice. It's much easier just to do the episodic care and not to worry about any of the future planning and so on because it's time consuming. So time would be a barrier." (Q03)

"I'd find a way to fund it so that I could say to GPs we're going to recognise your time, we're going to pay the opportunity cost of you not seeing your patients for a couple of hours and then I'd find a way for those patients that are not being seen to be seen by somebody somehow some other time. " (Q01)

Facilitators to implementation: Computer record system prompts

"That would need to have some sort of automatic prompt…actually integrated into the medical [records] system... everyone above the age of 75 would come up as prompter, you want to do this, quality of life, end of life assessment, yes/no." (Q03)

Incorporation into routine aged care assessment

" I think it would help to systematize the way we manage older people, so it could be something that you slotted in, you may not do it with the patient but might sit it somewhere alongside the Over 75 Health Assessment[1] because it's got lots of questions… to include falls and fractures, hospitalizations, nutrition, how people are managing at home, et cetera. And so that would fit quite nicely alongside and we would be able to complete that in parallel." (Q02)

1. The Over 75 Health assessment is a funded general review of patient wellbeing conducted for people over 75 years old, focusing on issues which may not come up in episodic consultations, like diet, social isolation or cognitive impairment.

CONSORT statement

Item	Page
1a Title – includes "randomised controlled trial"	1
1b Abstract. Reports Trial design, participant eligibility, participants, objective, outcomes, total number randomized but not by group, outcomes, conclusion. There is a strict word limit of 250 words, so other issues were not in the abstract: How randomisation occurred, recruitment/ trial status, numbers analysed in each group, Harms (not relevant), trial registration and funding (both reported elsewhere in the submission.	6-7
2a Background outlines rationale for study	10-12
2b Objectives – present	12
3a- trial design	13
3b – there were no changes to the original trial design	n/a
4a- participant selection described	14
4b- study setting described	13-14
5. Intervention described fully.	15-16
6a Outcomes described fully and how they were assessed described,	13
6b There were no changes to outcomes used.	n/a
7a - No sample size was calculated. This was a pilot trial. We reported the calculated Intracluster correlation and Design effect to assist investigators plan trials in the future is reported.	22
7b- Interim analysis and stopping guidelines – not relevant to this study	n/a
8a Randomisation sequence generation described	14
8b Type of randomisation described and blocking and stratification described	14
9 Randomisation allocation concealment described (central randomisation and not advising the participants which arm they were in)	14,16
10 Randomisation sequence generation and assignment described.	14
11a Blinding technique described	16
11b Description of interventions and data collection described, and interventions similar as much as possible	15-16
12a Analytical methods fully described.	17
12b There were no subgroup analyses.	n/a
13a Participant flow described and shown in Figure 1.	Figure 1
13b Losses and exclusions shown and described as above	19, Figure 1
14a Recruitment dates shown	19
14b Reason trial stopped was as per funding timeline. Not mentioned	n/a
15 Baseline data for each group shown (tables 1 for GPs and 2 for patients)	19 Tables 1,2
16 Analysis by assigned group shown	20-21, Tables 3-5
17a Outcomes with 95% confidence intervals shown – Tables 3-5	20-21, Tables

	3-5
17b Binary outcomes (died/ not died; predicted death/not predicted described, but absolute and relative risk ratios are not relevant.	n/a
18 No ancillary analyses performed.	n/a
19 Harms – not relevant to this study	n/a
20 Limitations discussed	25
21 Generalisability- As this is a pilot RCT, generisability is not a factor. We have included a cautionary statement in the strengths and weaknesses section.	24
22. Interpretation is consistent with results	23-24
23 Trial registration number reported in abstract and text	7,18
24 Trial protocol and data can be accessed	yes
25 Trial funding sources presented.	5